2009-06-10

Information Hiding by Stochastic Diffusion and its Application to Printed Document Authentication

Jonathan Blackledge
Technological University Dublin, jonathan.blackledge@tudublin.ie

Eugene Coyle
Technological University Dublin, Eugene.Coyle@tudublin.ie

Information Hiding by Stochastic Diffusion and its Application to Printed Document Authentication

Jonathan Blackledge[†] and Eugene Coyle[*]

SFI Stokes Professor of DSP
Faculty of Engineering
Dublin Institute of Technology

School of Electrical Engineering Systems
Faculty of Engineering
Dublin Institute of Technology

E-mail: [†]jonathan.blackledge@dit.ie [*]eugene.coyle@dit.ie

Abstract — **The use of image based information exchange has grown rapidly over the years in terms of both e-to-e image storage and transmission and in terms of maintaining paper documents in electronic form. Further, with the dramatic improvements in the quality of COTS (Commercial-Off-The-Shelf) printing and scanning devices, the ability to counterfeit electronic and printed documents has become a widespread problem. Consequently, there has been an increasing demand to develop digital watermarking, information hiding and covert encryption methods which can be applied to both electronic and printed images (and documents) for the purposes of authentication, prevent unauthorized copying and, in the case of printed documents, withstand abuse and degradation before and during scanning. In this paper we consider the background to a new method to hiding image based information by diffusing it with a stochastic field (uniformly distributed noise). This 'diffusion only' approach is used specifically to design a system for authenticating printed information that is robust to a low resolution 'print-scan cycle'.**

Keywords — **Information Hiding, Stochastic Diffusion, Print Authentication.**

I Introduction

In this paper, an approach to image information hiding is presented and some possible applications considered. It is based on computing a 'scrambled image' by diffusing an 'input image' with a stochastic field (a cipher). For e-to-e applications, a cover image (covertext) can be applied subject to a user defined diffusion-to-covertext ratio. The information is subsequently recovered by removing the covertext and then correlating the output with the original (key dependent) cipher. This approach provides the user with a method of hiding image-based information in a host image before transmission of the data. In this sense, the method provides a steganographic approach to transmitting encrypted information that is not apparent during an intercept [1]. Decryption is based on knowledge of the key(s) and access to the host image.

With regard to digital image analysis and e-to-e communications, the method provides a way of embedding information in an image that can be used for authentication from an identifiable source, a method that is relatively insensitive to lossy compression, making it well suited to digital image transmission. However, with regard to document authentication, use of a covertext is not robust. The reason for this is that the registration of pixels associated with a covertext can not be assured when the composite image is printed and scanned. We therefore consider a diffusion only approach to document authentication. This is because the process of diffusion (i.e. the convolution of information with a cipher) is compatible with the physical principles of an imaging systems and thus, with image capture devices (digital cameras and scanners, for example) that, by default, conform to the 'physics' of optical image formation [2].

The diffusion of plaintext (in this case, an image) with a stochastic field (the cipher) has a synergy with the encryption of plaintext using a cipher and an XOR operation (when both the plaintext and cipher are represented as binary streams) [3], [4]. However, decryption of a convolved im-

age (deconvolution) is not as simple as XORing the ciphertext with the appropriate cipher. Here, we consider an approach which is based on preconditioning the original cipher in such a way that decryption (de-diffusion) can be undertaken by correlating the ciphertext with the cipher. The output ciphertexts generated for printed document authentication are textures of a type that are determined by the spectral characteristics of the plaintext which can be applied using low resolution Commercial-Off-The-Shelf (COTS) printers and scanners. In this sense, the approach is based on a form of 'texture coding'.

II Stochastic Diffusion and Confusion

In terms of plaintexts, diffusion is concerned with the issue that, at least on a statistical basis, similar plaintexts should result in completely different ciphertexts even when encrypted with the same key. This requires that any element of the input block influences every element of the output block in an irregular fashion. In terms of a key, diffusion ensures that similar keys result in completely different ciphertexts even when used for encrypting the same block of plaintext. This requires that any element of the input should influence every element of the output in an irregular way. This property must also be valid for the decryption process because otherwise an intruder may be able to recover parts of the input from an observed output by a partly correct guess of the key used for encryption. The diffusion process is a function of the sensitivity to initial conditions, conditions that all cryptographic systems should have. Further, all cryptographic systems should exhibit an inherent topological transitivity causing the plaintext to be mixed through the action of the encryption process.

The process of 'confusion' ensures that the (statistical) properties of plaintext blocks are not reflected in the corresponding ciphertext blocks. Every ciphertext must have a random appearance to any observer and be quantifiable through appropriate statistical tests. Diffusion and confusion are processes that are of fundamental importance in the design and analysis of cryptological systems, not only for the encryption of plaintexts but for data transformation in general.

Consider the fundamental imaging equation given by [6]

$$u(x,y) = p(x,y) \otimes_2 u_0(x,y) + n(x,y)$$

where u_0 is the 'input' (information associated with the 'object plane') u is the output (information associated with the 'image plane'), p is the Point Spread Function (PSF), n is the noise function (where $\Pr[n(x,y)]$ is, ideally, known *a priori*[1]) and $\otimes_2$ denotes the two-dimensional convolution integral. In optics, both the operator $\otimes_2$ and the functional form of p are derived from solving a physical problem (using a Green's function solution) compounded in a particular Partial Differential Equation (e.g. the wave equation or the diffusion equation) [5]. For example, if u is taken to be due to the diffusion of light through an optical diffuser and thereby a solution to the diffusion equation (with initial condition u_0), then at a time T

$$p(x,y) = \exp\left(\frac{x^2+y^2}{4DT}\right)$$

where D is the 'Diffusivity' of the diffuser. This is an example of 'Gaussian diffusion' since the characteristic Point Spread Function is a Gaussian function. However, in general, a variety of PSFs are possible with regard to the fundamental imaging equation and the imaging systems to which it applies as a basic model. The PSF of an imaging system is fundamental to its characterisation and may be derived theoretically and/or experimentally. However, in Cryptology, we are 'free' to choose any PSF. Stochastic diffusion involves interchanging the roles of p and n, i.e. replacing $p(x,y)$ - a deterministic PSF - with $n(x,y)$ - a stochastic function. Thus, stochastic or 'noise' diffusion is compounded in the result

$$u(x,y) = n(x,y) \otimes_2 u_0(x,y) + p(x,y)$$

where p can now be any function including a stochastic function, i.e.

$$u(x,y) = n_1(x,y) \otimes_2 u_0(x,y) + n_2(x,y)$$

where both n_1 and n_2 are stochastic functions which may be of the same type (i.e. have the same PDFs) or of different types (with different PDFs).

The simplest form of stochastic diffusion is based on the equation

$$u(x,y) = n(x,y) \otimes_2 u_0(x,y).$$

There are two approaches to solving the inverse problem: Given u and n, obtain u_0. We can deconvolve by using the convolution theorem giving

$$u_0(x,y) = \mathcal{F}_2^{-1}\left[\frac{U(k_x,k_y)N^*(k_x,k_y)}{\mid N(k_x,k_y)\mid^2}\right]$$

where N is the Fourier transform of n, U is the Fourier transform of u, $\mathcal{F}_2^{-1}$ denotes the (two-dimensional) inverse Fourier transform and k_x and k_y are the spatial frequencies in the Fourier plane. However, this approach requires regularization in order to eliminate any singularities when $\mid N \mid^2 \rightarrow 0$ through application of a constrained deconvolution filter such as the Wiener filter [7]. Alternatively, if n is the result of some random number generating algorithm, we can construct the

[1] Pr denotes the Probability Density Function.

stochastic field

$$m(x,y) = \mathcal{F}_2^{-1}\left[\frac{N^*(k_x,k_y)}{\mid N(k_x,k_y)\mid^2}\right]$$

where $\mid N(k_x,k_y)\mid^2 > 0$, the diffused field now being given by

$$u(x,y) = m(x,y) \otimes_2 u_0(x,y).$$

The inverse problem is then solved by correlating (denoted by $\odot_2$) u with n, since[2]

$$n(x,y) \odot_2 u(x,y) \Longleftrightarrow N^*(k_x,k_y)U(k_x,k_y)$$

and

$$N^*(k_x,k_y)U(k_x,k_y)$$
$$= N^*(k_x,k_y)M(k_x,k_y)U_0(k_x,k_y)$$
$$= N^*(k_x,k_y)\frac{N^*(k_x,k_y)}{\mid N(k_x,k_y)\mid^2}U_0(k_x,k_y) = U_0(k_x,k_y)$$

so that

$$u_0(x,y) = n(x,y) \odot_2 u(x,y).$$

The condition that $\mid N(k_x,k_y)\mid^2 > 0$ is simply achieved by implementing the following process: $\forall k_x, k_y$, if $\mid N(k_x,k_y)\mid^2 = 0$, then $\mid N(k_x,k_y)\mid^2 = 1$. This result can be used to 'hide' an image in another image as discussed in the following section.

III Digital Image Watermarking

Consider the case when we have two independent images $i_1(x,y) \geq 0 \forall x,y$ and $i_2(x,y) \geq 0 \forall x,y$ and we consider the case of embedding i_1 with i_2. We construct a stochastic field $m(x,y) \geq 0 \forall x,y$ *a priori* and consider the equation

$$u(x,y) = rm(x,y) \otimes_2 i_1(x,y) + i_2(x,y) \qquad (1)$$

where

$$\|m(x,y) \otimes_2 i_1(x,y)\|_\infty = 1 \quad \text{and} \quad \|i_2(x,y)\|_\infty = 1.$$

By normalising the terms in this way, the coefficient $0 \leq r \leq 1$ can be used to adjust the relative magnitudes of the terms such that the diffused image i_1 is a perturbation of the 'host image' (covertext) i_2. This provides us with a way of digital watermarking [8] one image with another, r being referred to as the watermarking ratio[3]. For applications in image watermarking, stochastic diffusion has two principal advantages: (i) a stochastic field provides more uniform diffusion than a deterministic function does; (ii) stochastic fields can be generated using random number generators that depend on a single initial value or seed (i.e. a private key). An example of this approach is shown in Figure 1. Here, an image i_2 (the 'host image') is watermarked with another image i_1 using stochastic diffusion. Because $r = 0.1$, the output u is only slightly perturbed by the stochastic field $m(x,y) \otimes_2 i_1(x,y)$ and hence $u \simeq i_2$ (at least from a visual perspective).

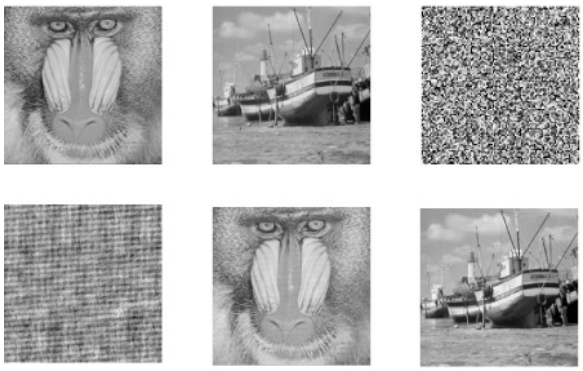

Fig. 1: Example of watermarking one image with another using stochastic diffusion. The 'host image' i_2 (top-left) is watermarked with the 'watermark image' i_1 (top-centre) using the diffuser (top-right) which is uniformly distributed random noise n whose pixel-by-pixel values depend upon the seed used (the private key). The result of computing $m \otimes_2 i_1$ (bottom-left) is added to the host image for $r = 0.1$ in equation (1) to generate the watermarked image u (bottom-centre). Recovery of the watermark image i_1 (bottom-right) is accomplished by subtracting the host image from the watermarked image and correlating the result with n.

a) Stochastic Field Generation

The stochastic field n used to compute m, can be generated using a range of (uniformly distributed) pseudo random number generators based on conventional (substitution) encryption algorithms coupled with existing key exchange protocols. The output array $\mathbf{n}$ is normalized so that $\|\mathbf{n}\|_\infty = 1$ and used to generate $n(x,y)$ on a row-by-row or column-by-column basis. Recovery of the watermark image requires knowledge of two keys: (i) the key used to generate n; (ii) the host image i_2.

b) Statistical Analysis

The expected statistical distribution associated with stochastic diffusion is Gaussian. This can be shown if we consider i_1 to be a strictly deterministic function described by a sum of N delta functions. Thus if

$$i_1(x,y) = \sum_{i=1}^{N}\sum_{j=1}^{N}\delta(x-x_i)\delta(y-y_j)$$

then

$$u(x,y) = m(x,y) \otimes_2 i(x,y)$$

[2]Where $\Longleftrightarrow$ denotes the transformation from image to Fourier space.

[3]Equivalent, in this application, to the standard term 'Signal-to-Noise' or SNR ratio as used in signal and image analysis.

$$= \sum_{i=1}^{N}\sum_{j=1}^{N} m(x-x_i, y-y_j).$$

Each function $m(x-x_i, y-y_j)$ is just $m(x,y)$ shifted by x_i, y_j and will thus be identically distributed. Hence, from the Central Limit Theorem

$$\Pr[u(x,y)] = \Pr\left[\sum_{i=1}^{N}\sum_{j=1}^{N} m(x-x_i, y-y_j)\right] =$$

$$\prod_{i=1}^{N}{}_{\otimes} \Pr[m(x,y)] \equiv \Pr[m(x,y)] \otimes_2 \Pr[m(x,y)] \otimes_2 ...$$

and we can expect $\Pr[u(x,y)]$ to be normally distributed for large N.

IV Print Authentication

The approach discussed in Section III can be used generally for e-to-e type communications where there is no loss of information. Steganography and watermarking techniques for hardcopy data (print) authentication have to be robust to the significant distortions generated by the printing and scanning processes as well as general soiling incurred through day-to-day use. If a watermarked image is printed and scanned back into electronic form, then the print/scan process will yield an array of pixels that will be different from the original electronic image even though it might 'look' similar. These differences can include the size of the image, its orientation, brightness, contrast and so on.

a) Diffusion Based Method

With respect to equation (1), of all the processes involved in the recovery of the watermark, the subtraction of the host image from the watermarked image is critical. If this process is not accurate on a pixel-by-pixel basis and deregistered for any of many reasons, then recovery of the watermark by correlation will not be effective. However, if we make use of the diffusion process alone, then the watermark can be recovered via a print/scan because of the compatibility of the optical processes involved (i.e. convolution of an object function with the PSF). Depending on the printing process applied, a number of distortions will occur which diffuse the information being printed. Thus, in general, we can consider the printing process to introduce an effect that can be represented by the convolution equation

$$u_{\text{print}} = p_{\text{print}} \otimes_2 u.$$

where u is the original electronic form of a diffused image (i.e. $u = m \otimes_2 i$ where i is the input image) and p_{print} is the PSF of the printer. An incoherent image of the data, obtained using a flat bed scanner, for example (or any other incoherent optical imaging system), will also have a characteristic PSF p_{scan}. Thus, we can consider a scanned image to be given by

$$u_{\text{scan}} = p_{\text{scan}} \otimes_2 u_{\text{print}}$$

where u_{scan} is taken to be the digital image obtained from the scan. Now, because convolution is commutative, we can write

$$u_{\text{scan}} = p_{\text{scan/print}} \otimes_2 m \otimes_2 i$$

$$= m \otimes_2 p_{\text{scan/print}} \otimes_2 i$$

where

$$p_{\text{scan/print}} = p_{\text{scan}} \otimes_2 p_{\text{print}}$$

which is the print/scan point spread function associated with the processing cycle of printing the image and then scanning it. Thus, the process $u(x,y) = m(x,y) \otimes_2 i(x,y)$ used to generate the data u and the process $i(x,y) = n(x,y) \odot_2 u(x,y)$ used to recover the image i produces a reconstruction for i whose fidelity is determined by the scan/print PSF. The principal requirement to do this in practice, is to re-size the scanned image back to the size of the original digital image i. This is due to the scaling relationship (for a function f with Fourier transform F)

$$f(\alpha x, \beta y) \Longleftrightarrow \frac{1}{\alpha\beta} F\left(\frac{k_x}{\alpha}, \frac{k_y}{\beta}\right).$$

The size of an image captured by a scanner or other device will depend on the resolution used. The size of the image obtained will inevitably be different from the original because of the resolution and window size used to print the diffused image u and the resolution used to scan the image. Since scaling in the spatial domain causes inverse scaling in the Fourier domain, the scaling effect must be 'inverted' before the watermark can be recovered by correlation since correlation is not a scale invariant process. Re-sizing the image (using an appropriate interpolation scheme such as the bi-cubic method, for example) requires a set of two numbers N_1 and N_2 (i.e. the $N_1 \times N_2$ array used to generate n and m) that, along with the key required to regenerate n, provides the 'private keys' needed to recover the data from the diffused image.

An example of this approach is given in Figure 2 which shows the result of reconstructing four different images (a photograph, finger-print, signature and text) used in the design of an impersonalized bank card. The use of 'diffusion only' watermarking for print security can be undertaken in colour by applying exactly the same diffusion/reconstruction methods to the red, green and blue components independently (as illustrated in Figure 2). Because this method is based on convolution alone, the reconstruction is not negated

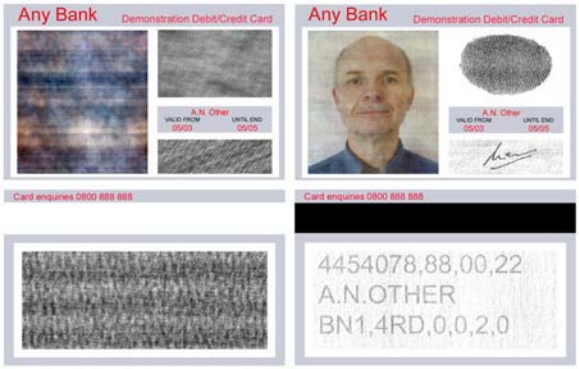

Fig. 2: Example of the application of 'diffusion only' watermarking. In this example, four images of a face, fingerprint, signature and text have been diffused using the same stochastic field m and printed on the front (top-left) and back (bottom-left) of an impersonalized identity card using a 600 dpi printer. The reconstructions (top-right and bottom-right, respectively) are obtained using a conventional flat-bed scanner operating at 300 dpi.

by any distortion of the PSF associated with the print/scan process, just limited or otherwise by its characteristics. Thus, if an image is obtained of the printed data field which is out of focus due to the characteristics of $p_{\text{scan/print}}$, then the reconstruction of will be out of focus to the same degree. Decryption of images with this characteristic is only possible using an encryption scheme that is based on a 'diffusion only' approach. The tolerance of this method to printing and scanning is excellent (details of which lie beyond the scope and extent of this paper) provided the output is cropped accurately (to within a few pixels) and oriented correctly.

Figure 3 shows another example of the technique applied to a composite image scanned from a passport, an application which is cheap and simple to implement with regard to authenticating a passport holders personal information. The degradation associated with the reconstruction is due to the low resolution of the printing and scanning rather than the information hiding method. Unless the correct stochastic field is used (as determined by the keys), it is not possible to reconstruct the image making counterfeiting or forgery improbable.

b) Covert Information Hiding

Digital watermarking is usually considered to be a method in which the watermark is embedded into a host image in an unobtrusive or near-unobtrusive way. In the context of the approach considered here, this can be achieved if we diffuse the host image with another image to generate a stochastic field.

Consider two images i_1 and i_2. Suppose we con-

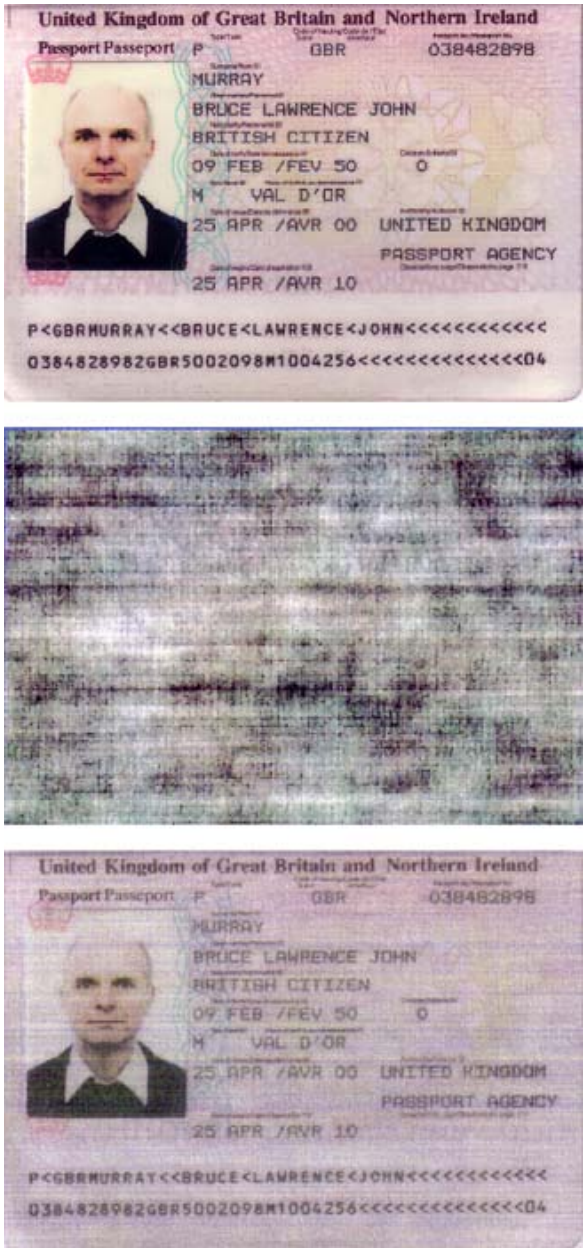

Fig. 3: Example of the stochastic diffusion method applied to passport authentication: Original image scanned from a passport at 400dpi (above), printed image after applying stochastic diffusion (centre) and reconstruction after scanning the printed stochastic field at 300dpi (below).

struct the following function

$$n = \mathcal{F}_2^{-1}\left(\frac{I_1}{\mid I_1 \mid^2} I_2\right)$$

where $I_1 = \mathcal{F}_2[i_1]$ and $I_2 = \mathcal{F}_2[i_2]$. If we now correlate n with i_1, then from the correlation theorem

$$i_1 \odot_2 n \Longleftrightarrow I_1^* \frac{I_1}{\mid I_1 \mid^2} I_2 \Longleftrightarrow i_2.$$

In other words, we can recover i_2 from i_1 with knowledge of n. Because this process is based on convolution and correlation alone, it is compatible and robust to printing and scanning, i.e. incoherent optical imaging as discussed in Section IV(a). An example of this is given in Figure 4. In this scheme, the host image can be considered to be a 'public key' and the stochastic field n the 'private key' required to reconstruct the 'hidden image'. Clearly, in the context of this public/private key paradigm, the 'private key' needs to be encrypted in order to ensure the security of any system that is based on this approach.

V Application to Covert Encryption

One of the principal components associated with the development of methods and algorithms to 'break' ciphertext is the analysis of the output generated by an attempted decrypt and its evaluation in terms of an expected type. The output type is normally assumed to be plaintext, i.e. to be in the form of characters, words and phrases associated with a natural language. For e-to-e applications, if a plaintext document is converted into an image file, then the method described in Section IV(b) on 'Covert Information Hiding' can be used to diffuse the plaintext image i_2 using any other image i_1 to produce the field n. If both i_1 and n are then encrypted, any attack on these data will not be able to make use of an 'analysis cycle' which is based on the assumption that the decrypted output is plaintext. This approach provides the user with a relatively simple method of 'confusing' the cryptanalyst and invalidates attack strategies that have been designed and developed on the assumption that the encrypted data have been derived from plaintext alone.

VI Conclusions

The approach discussed in this paper is robust to a wide variety of attacks including geometric attacks, drawing, crumpling and print/scan attacks, details of which lie beyond the scope of the paper. The method is relatively insensitive to lossy compression, filtering, amplitude adjustments, additive noise and thresholding (for a binary input). The principal weakness of the system is its sensitivity to rotation and cropping. This can be minimized by orienting the document correctly and accurately before scanning and using automatic cropping software which is available with selected scanners (e.g. Cannon scanners).

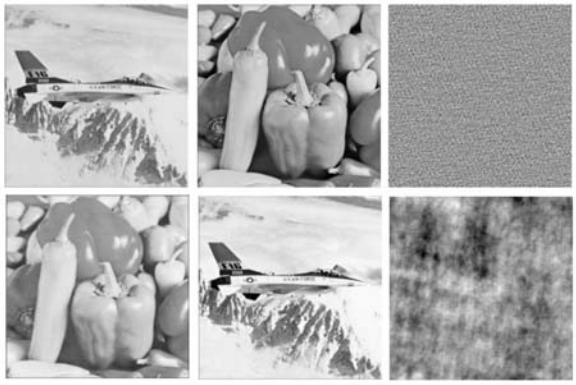

Fig. 4: Example of a covert watermarking scheme. i_1 (top-left) is 'processed' with i_2 (top-middle) to produce the stochastic field n (top-right). i_2 is printed at 600 dpi, scanned at 300 dpi and then re-sampled back to its original size (bottom-left). Correlating this image with n generates the reconstruction (bottom-centre). The reconstruction depends on just the host image and n. If the n and/or the host image are different or corrupted, then a reconstruction is not achieved, as in the example given (bottom-right).

The 'visibility' of the ciphertext generated through the process of stochastic diffusion (e.g. Figures 2 and 3) and the compatibility of this approach with the physical principles of an imaging system (convolution of an object function with the PSF), increase the robustness associated with the retrieval of the 'hidden information' after scanning at low resolution.

References

[1] N. Ferguson and B. Schneier B, "Practical Cryptography", *Wiley*, 2003.

[2] M. Born and E. Wolf, "Principles of Optics (6th Edition)", *Pergamon Press*, Oxford, 1980.

[3] J. Buchmann, "Introduction to Cryptography", *Springer*, 2001.

[4] O. Goldreich, "Foundations of Cryptography", *Cambridge University Press*, 2001.

[5] E. G. Steward, "Fourier Optics: An Introduction", *Horwood Scientific Publishing*, 1987.

[6] J. M. Blackledge, "Digital Image Processing", *Horwood Publishing*, 2005.

[7] M. Bertero and B. Boccacci, "Introduction to Inverse Problems in Imaging", *Institute of Physics Publishing*, 1998.

[8] I. J. Cox, M. L. Miller and J. A. Bloom, "Digital Watermarking", *Morgan Kaufmann*, 2002.